



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/595,719	06/16/2000	Minos N. Garofalakis	Garofalakis-6-1-36-11-10	1456

7590 03/05/2004

Stephen J Weed Esquire
Synnestvedt & Lechner LLP
2600 Aramark tower
1101 Market Street
Philadelphia, PA 19107-2950

EXAMINER

HUYNH, CONG LAC T

ART UNIT	PAPER NUMBER
----------	--------------

2178

DATE MAILED: 03/05/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/595,719

Applicant(s)

GAROFALAKIS ET AL.

Examiner

Cong-Lac Huynh

Art Unit

2178

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 16 June 2000.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-22 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1,2,7,8 and 13-22 is/are rejected.
- 7) ☒ Claim(s) 3-6 and 9-12 is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 16 June 2000 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date 3.
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: _____.

DETAILED ACTION

1. This action is responsive to communications: the application filed on 6/16/00, and the IDS filed on 9/18/00.
2. Claims 1-22 are pending in the case. Claims 1, 7, 14, 16, 18 are independent claims.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).
5. Claims 1-2, 7-8, 13-22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Tateno (US Pat No. 5,812,999, 9/22/98, filed 3/13/96).

Regarding independent claim 1, Tateno discloses:

- generalizing input sequences associated with a document to develop general sequences, said input sequences reflecting the structure of a document (figures 4-5: the sequence as in figure 4 reflect the structure of the document)
- selecting a document descriptor from said input sequences, said general sequences (col 3, lines 6-63: DTD 40 is selected from the input sequence that reflects the document structure)

Tateno does not explicitly disclose factoring said input sequences and said general sequences to develop factored sequences where said factored sequences use minimum descriptor length (MDL) principles.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Tateno to include factoring the input sequence and said general sequences to develop factored sequences since the input sequence as in Tateno, if including two or more *repeated elements*, then said elements can be factored the same way as factoring in the multiplication to eliminate the repeated numbers in multiplying to obtain a *shorter sequence* of the same value but including all of the actual elements of the sequence.

In addition, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Tateno to include using the minimum descriptor length principles for the factored sequence since as just mentioned, factoring a sequence provides a *shorter sequence*. This suggests that performing factoring an

Art Unit: 2178

input sequence many times would provide a shortest sequence that reflects the minimum length of the DTD of the document.

Regarding claim 2, which is dependent on claim 1, Tateno suggests selecting a document descriptor which encompasses all of said input sequences and exhibits a minimum MDL cost as mentioned in claim 1 above.

Tateno does not explicitly disclose encoding said input sequences, said general sequences, and said factored sequences.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Tateno to include the step of encoding into Tateno since encoding documents is an obvious step in programming process.

Claims 7-8, 13 include the same limitations as in claims 1-2, and are rejected under the same rationale.

Regarding independent claim 16, Tateno discloses:

- discovering OR patterns among said input sequences (col 2, lines 49-65: symbol "[]" is the OR pattern in the input sequence "(title, (paragraph|figure)*, chapter*)")

Tateno does not explicitly discovering sequence patterns among said input sequences and OR patterns. However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Tateno to include discovering sequence patterns among said input sequences and OR patterns since any element in

Art Unit: 2178

the sequence as in figures 4-5 such as "paragraph" or "chapter" can be used as the sequence patterns.

Regarding claim 17, which is dependent on claim 16, Tateno discloses that discovering OR patterns comprises partitioning said input sequences (col 5, lines 15-46: input sequence of the document is partitioned and divided into the reference units).

Regarding independent claim 18, Tateno discloses:

Generalizing input sequences, said generalizing comprises:

- discovering OR patterns among said input sequences (col 2, lines 49-65: symbol "[]" is the OR pattern in the input sequence "(title, (paragraph|figure)*, chapter*)")

Selecting a document descriptor from said input sequence and said general sequences (col 3, lines 6-40: selecting the DTD of the tag sequence as in figure 4).

Tateno does not explicitly discovering sequence patterns among said input sequences and OR patterns. However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Tateno to include discovering sequence patterns among said input sequences and OR patterns since any element in the sequence as in figures 4-5 such as "paragraph" or "chapter" can be used as the sequence patterns.

Art Unit: 2178

Regarding claim 19, which is dependent on claim 18, Tateno discloses that discovering OR patterns comprises partitioning said input sequences (col 5, lines 15-46: input sequence of the document is partitioned and divided into the reference units).

Regarding claim 20, which is dependent on claim 19, Tateno does not disclose explicitly that factoring said input sequences and said general sequences to develop factored sequences, wherein said factored sequences are available to said selecting.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Tateno to include the factoring step into Tateno since the input sequence as in Tateno, if including two or more *repeated elements*, then said elements can be factored the same way as factoring in the multiplication to eliminate the repeated numbers in multiplying to obtain a shorter sequence of the same value but including all of the actual elements of the sequence.

Regarding claim 21, which is dependent on claim 20, Tateno does not disclose explicitly that said selecting employs minimum descriptor length (MDL) principles.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have include the minimum descriptor length principles into Tateno since as mentioned in claim 20 above, factoring a sequence provides a shorter sequence. This suggests that performing factoring an input sequence many times would provide a shortest sequence that reflects the minimum length of the DTD of the document.

Art Unit: 2178

Regarding claim 22, which is dependent on claim 21, Tateno discloses that said document descriptor is a document type descriptor (DTD) and said document is a SGML document (col 3, lines 31-48; col 1, lines 34-56; col 2, lines 6-41).

Tateno does not disclose that said document is an extensible Markup Language (XML) document. However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Tateno to include the XML document into Tateno since it was well known that XML is a slimmed-down version of SGML.

Claims 14-15 are for a computer readable medium of method claims 16-17, 18-19, and are rejected under the same rationale.

Allowable Subject Matter

6. Claims 3-6, 9-12 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

Conclusion

7. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Beaverson et al. (US Pat No. 5,299,206, 3/29/94, filed 10/24/91).

Kuwahara (US Pat No. 6,202,072 B1, 3/13/01, filed 12/5/97).

Sundaresan (US Pat No. 6,569,207 B1, 5/27/03, filed 10/5/98).

Art Unit: 2178

Sundaresan (US Pat No. 6,487,566 B1, 11/26/02, filed 10/5/98).

Ting (US Pat No. 5,930,746, 7/27/99, filed 8/9/96).

Nakao (US Pat No. 6,061,697, 5/9/00, filed 8/25/97).

Nasr et al. (US Pat No. 6,438,540 B2, 8/20/02, filed 6/19/01).

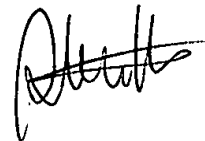
Chen et al. (US Pat No. 6,507,856 B1, 1/14/03, filed 1/5/99).

8. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Cong-Lac Huynh whose telephone number is 703-305-0432. The examiner can normally be reached on Mon-Fri (8:30-6:00).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on 703-308-5186. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Clh
3/1/04



**STEPHEN S. HONG
PRIMARY EXAMINER**